图书情报工作
第 63 卷 第 17 期 2019 年 9 月

国内外科学数据开放共享研究综述*

■ 盛小平[1] 武彤[2]

[1] 上海大学图书情报与档案系 上海 200444 [2] 华南师范大学经济与管理学院 广州 510006

摘要：[**目的/意义**]通过综述国内外科学数据开放共享研究成果，为促进国内相关理论研究与实践发展提供参考。[**方法/过程**]运用文本内容分析法，从科学数据开放共享政策、平台与项目、利益与障碍、机制与模式、利益相关者角色与责任、应用与调查、对策与建议 7 个方面解析国内外研究进展。[**结果/结论**]目前国内外对科学数据开放共享主题进行了广泛探索，特别是在科学数据开放共享政策与平台、机制与模式、利益与障碍等关键问题上有了一些新认识。但整体上仍处于起步阶段，仍需从有效区分科学数据开放共享与科学数据共享的概念与行为、建立健全科学数据开放共享机制和加强科学数据开放共享实践调查分析等方面提升科学数据开放共享研究水平。

关键词：科学数据 开放共享 研究综述

分类号：G203

DOI：10.13266/j.issn.0252－3116.2019.17.001

科学数据开放共享可追溯到 20 世纪 50 年代，在当时分享国际科学理事会（International Council of Science）1957－1958 年收集的数据时，出现了科学数据开放获取概念。然而，随后对开放获取科学数据的研究要比开放获取学术出版物少得多[1]。如今，科学数据开放共享已经引起人们的重视。一方面，开放获取科学数据被视为与所有学科领域的各种数据和所有研究人员相关[2]；另一方面，国内外颁布了许多科学数据开放共享政策，启动了许多科学数据开放共享计划或项目，同时发表了一些相关研究成果。2019 年 1 月 15 日，笔者首先以“科学数据”或“科研数据”“研究数据”为篇名，并把“开放共享”或“开放获取”“开放存取”作为并列篇名，在中国知网中进行“精确”检索后得到 20 篇文献；然后，以同样方式在国家图书馆馆藏书目中进行“组合检索”后发现有 5 条记录；同时以“scientific data”或“research data”作为“标题（title）”检索词，与“open sharing”或“open access”作为“关键词”或“主题词”，组合检索国外数据库（包括 Elsevier、Springer、EBSCO、Emerald、Taylor & Francis、SAGE Premier、Wiley、Proquest），得到 65 篇相关文献；另外利用 Dogpile 搜索引擎，输入“open sharing of research data”或“open access to research data”，检索互联网后得到 15 篇相关论文与研究报告。虽然有文献分别综述了国外开放科学数据研究[3－4]和科学数据共享研究[5－6]，但是科学数据开放共享研究主题没有得到有效揭示。为此，本文从科学数据开放共享政策、平台与项目、利益与障碍、机制与模式、利益相关者角色与责任、应用与调查、对策与建议 7 个方面来综述国内外研究成果。

1 对科学数据开放共享政策的研究

科学数据开放共享政策成为研究焦点之一。许多国际组织如世界经合组织、欧盟委员会、欧洲研究理事会、欧洲研究型大学联盟、科学国际联盟、国际北极科学委员会、国际科学技术数据委员会、国际科学理事会世界数据系统、国际地球观测组、八国集团，和美国、英国、加拿大、德国、澳大利亚、瑞典、挪威、爱尔兰、芬兰、丹麦、荷兰、斯洛文尼亚、印度等多个国家政府机构、研究机构、研究资助机构、出版商等利益相关者都制定了科学数据开放共享政策[7－11]。这些科学数据开放共享政策通常可分为宏观、中观和微观 3 个层面[12]。

政府、科研资助机构、科研教育机构、学术出版机构制定的科学数据开放共享政策在政策要素或内容及

* 本文系国家社会科学基金项目“开放科学环境下的科学数据开放共享机制与对策研究”（项目编号：18ATQ007）研究成果之一。

作者简介：盛小平（ORCID：0000－0002－6341－6973），教授，博士，博士生导师，E-mail：shengxp68@126.com；武彤（ORCID：0000－0001－8514－4684），硕士研究生。

收稿日期：2018－10－11 修回日期：2019－04－02 本文起止页码：6－14 本文责任编辑：王传清

盛小平，武彤. 国内外科学数据开放共享研究综述[J]. 图书情报工作,2019,63(17):6－14.

其实施方案上有许多差异,但这些政策举措主要围绕科学数据交流渠道各环节[13],集中体现在数据存储方式、数据权利与利益相关者责任、开放数据范围、开放获取方式、数据质量、数据互操作、数据开放时滞、知识产权保护、数据安全等,呈现出如下几个鲜明特征:注重科学数据开放共享的价值,注重科学数据的知识产权及其保护,注重科学数据的互操作,注重科学数据的高质量,注重科学数据的评估,强调利益相关者的责任担当[14]。

尽管有文献把科学数据开放共享政策与科学数据共享政策[15]或机构数据政策[16]混为一谈,但是,它们并非同一个概念。科学数据开放共享政策是指某机构或组织对如何有效实施科学数据开放共享的相关规定或制度安排,有时也名为科学数据开放获取政策。但它不同于传统的科学数据共享政策,后者往往要求在特定领域或有限范围内且遵循某些条件来实施科学数据的共享,仅适用于特定范围内的用户,而对其他用户是无法利用的[10]。

2 对科学数据开放共享平台与项目的研究

如今,国内外已经建立多个科学数据开放共享平台,这主要包括:①科学数据开放注册平台,如 DataCite、re3data.org、开放获取目录数据知识库[7]91、Dataverse 网络[17]、个人基因项目全球网络[18]等。②科学数据开放存储与服务平台,如 Dryad、Figshare、PANGAEA、GBIF[7]95-100,192-193、Neurotycho.org[19]、开放光谱数据库[20]、耶鲁大学社会和政策研究开放数据知识库[21]、蛋白质数据库[22]、北京大学开放研究数据平台[23]等。③科学数据开放出版平台,如生物医学中心(BioMed Central)、公共医学中心(PubMed Central)、《科学》期刊、《自然》期刊、PLoS[7]102 等。现有研究主要从数据知识库结构[24]、软件栈(software stack)设计[20]、技术基础设施[7]95-100,192-193、数据特性[19]、数据资源与服务[22]、平台建设内容和应用效果[23,25]、建设过程与挑战以及经验[21]等方面进行了广泛探索。其中,科学数据开放注册平台为个人或机构提交与存储相关领域的科学数据提供了通道;科学数据开放存储与服务平台可以在开源的基础上实施开放采用软件(open adoption software)解决方案[26],既可以储存科学数据,也可以提供数据论文的浏览与检索、数据链接甚至数据关联的分析;科学数据开放出版平台搭建了科学数据走向开放出版的桥梁,也成为科学数据开放出版的“集市”。从印度机构知识库来看,大多数机构知识库收藏的文献较少,其中期刊论文所占比例较高,会议论文与学位论文料所占比例较低,而预印本/工作论文、教学资源和专利的数量非常少[27]。

在开发平台的同时,国内外已经实施了许多科学数据开放共享项目。自 2007 年开始,欧盟先后资助了许多科学数据项目,如 RECODE、ARIANDE、SERSCIDA、ODE、EUROVO - AIDA、CLARIN、CESSDA、ELIXIR[28],以及 ACUMEN、OAPEN、OpenAIRE、OpenAIREplus、PEER、SOAP[29]等。这些项目多以科学数据共享和开放获取的基础设施建设为目标,旨在从不同角度促进科学数据的存储、共享和复用。例如,OpenAIRE 项目是为欧洲研究人员建立一个合作的开放获取基础设施,使他们能够访问开放获取信息,并支持建立欧洲层面上的开放获取政策和向决策者提供使用统计数据[30]。欧洲开放项目(European project OpenUp)在成员国家科学数据之间实现了科学数据的可比性、灵活性、一致性、可访问性、可重现性[31]。

3 对科学数据开放共享利益与障碍的研究

科学数据开放共享具有许多巨大的利益,这些利益也是利益相关者参与科学数据开放共享实践的动力。不过,目前,科学数据开放共享还面临许多障碍。

3.1 对科学数据开放共享利益的研究

国外科学数据开放共享的利益研究始于 20 世纪 70、80 年代[32]。整体而言,科学数据开放共享可以产生如下众多利益:①经济利益,这包括:可使科研人员在申请某些科研项目时能够获得科研资助机构的资助(因为开放共享研究数据是某些科研项目得到资助的前提条件之一);获得公共投资的更大回报[33];加快研究成果向产品和服务的转化,并使这些成果更快地进入市场[34];刺激下游应用和商业创新[35];获得更大的竞争力[36];可以促进经济增长,发挥科学数据的经济效益[37]。②政治与社会利益,这包括:有利于提高政府信誉,改善政府治理,并提高政府资助科学研究的透明度[38];有利于公众了解国家的科学研究[39];向公众提供公共资助的研究成果[40];促进公众参与科学研究,提高公众参与度和满意度[38];有助于向发展中国家转移事实性数据和促进其能力建设[35]。③技术和操作利益,这包括:有助于加强开放科学调查;鼓励分析和意见的多样性;发现新的研究主题,促进新的研究;再现和验证研究过程与结果,提升研究成果质量;

避免重复工作[41]；更广泛获取、使用和再利用科学数据；创建更完整和透明的科学记录[42]；提高科学透明度和准确性，产生更高的科学研究影响[43]；加速研究成果的转化与推广[44]；提升跨学科、跨部门和跨机构以及国际研究；有助于最大限度地发挥新的数字技术和网络的研究潜力[35]；有利于鼓励合作，改进对危机的响应[34]；可以提高论文以及科学数据的引用率[45]；可以提高研究机构和研究者的学术声誉和知名度[46]。总之，科学数据开放共享可以为利益相关者带来巨大的利益。但不同利益相关者得到的利益可能差异较大，需要建立利益平衡机制[47]。

3.2 对科学数据开放共享障碍的研究

科学数据开放共享障碍主要包括如下几个方面[36,38]：①机构本身存在的科学数据开放共享障碍，比如：缺乏分享数据的政策，存在风险规避文化，没有公布数据的统一政策，只公开非价值数据，没有有效宣传数据，靠数据创造收入，缺少激励措施等。②处理科学数据开放共享工作任务面临的障碍，比如：没有能力发现有价值的数据，无法访问原始数据，没有关于开放数据质量的信息，没有索引或其他手段确保容易搜索到正确的数据，用户不知道数据潜在用途，数据格式和数据集太复杂以至于难以处理和使用，没有支持工具或帮助桌面等，特别是在开放获取数字学位论文中的研究数据时仍面临巨大挑战[48]。③使用科学数据和参与科学数据开放共享面临的障碍，比如：不鼓励用户共享数据，不明确数据用户的需求，没有处理用户输入的流程，忽视开放数据机会，缺乏时间，在下载数据之前必须注册，担心错误的结论和潜在的负面后果，缺乏利用或理解数据的知识，缺乏使用信息的必要能力，机构之间缺乏合作资助途径，受商业合同的限制等。④与科学数据开放共享相关的法律障碍，比如：缺乏利用数据的法律框架，新法律带来的不确定性，存在相互矛盾的数据共享和使用权利规则或法律，存在隐私问题，缺乏开放许可，存在诉讼威胁，存在禁止商业使用等。⑤科学数据或信息质量存在的障碍，比如：缺乏信息的准确性，存在不完整的信息，存在过时和无效的数据，存在信息冗余，存在数据可用性问题[49]等。⑥与科学数据开放共享相关的技术障碍，比如：各机构之间格式不兼容，缺乏技术数据基础设施，无法利用众包，没有利用用户反馈的程序，缺乏元数据，缺乏机器可读的格式，使用开放数据门户复杂，缺乏编程和统计技能，存在语言障碍等。

我国科学数据开放共享的主要障碍是缺乏明确的国家层面的法律体系，已制定的相关政策并不具备法律效力；缺乏行之有效的管理体制；缺乏完善的数据归档、汇交和开放共享平台；数据质量和可获得性不高；存在信息封锁和信息壁垒现象；政府掌握的许多数据资源处于闲置状态；网上中文数据所占比例很小；科学数据市场调节基本处于失灵状态；缺乏数据资源开放共享领域的专业人才；数据生产和管理者不愿意或者不能够将数据开放共享，需进一步加强科研项目资助体系中的制度约束和政策支持[50-51]。

4 对科学数据开放共享机制与模式的研究

科学数据开放共享机制与模式也是开放科学研究领域的焦点问题。

4.1 对科学数据开放共享机制的研究

科学数据开放共享机制是指科学数据开放共享系统要素之间的相互作用过程、方式与运行规律，包含科学数据汇交机制、出版机制、传播机制、联盟机制、激励机制、开放共享评价机制、开放共享保障机制等。其中，科学数据汇交机制要求国家财政经费支持的行业部门和科研项目以及数据中心产生和管理的数据资源，按照国家《科学数据管理办法》和各部门制定的数据汇交管理办法，实施自下而上的汇交或交换；科学数据出版机制要求在互联网上公开数据，或者作者在发表论文前将科学数据存储在可公开访问的相关数据库中，支持除数据提供者之外的研究人员或者组织机构下载、分析、再利用以及引用数据[52]；科学数据传播机制需要增加搜索引擎收录量、加强用户对数据开放平台的使用、利用社交网络提高开放共享数据资源的传播速度与范围、提升数据资源的利用率，来最大限度地发挥科学数据的价值[52]；科学数据联盟机制促使数据（平台）中心之间遵循共同协商的原则或声明，利用统一的标准或接口（软件系统），进行元数据或数据的互操作[52]；科学数据激励机制需要采取有效的激励措施促使数据拥有者将手里的数据开放出来，最大限度发挥科学大数据的潜力与能量；科学数据开放共享评价机制需要在数据开放共享过程中，采取严格的同行评议来确保数据的质量、可用性、完整性、规范性，并在数据开放共享后，需要监管部门、资助机构等定期或不定期地对数据开放情况进行调查和监督，还要对开放数据的质量、可用性和价值以及影响力进行评估[53]；科学数据开放共享保障机制需要建立健全宏观、中观和微观的政策法规与制度以及技术平台，同时构建一个

跨领域的区域性组织与服务保障体系[12]。这些机制可以成为国内机构实施科学数据开放共享的战略选择与行动指南。

4.2 对科学数据开放共享模式的研究

科学数据开放共享模式涉及科学数据开放共享的运行模式与资助模式。从科学数据开放共享的公益性与商业性来看，一个国家或地区内的科学数据开放共享模式通常有两种，即政府主导下的公益性运行模式和市场主导下的商业化运行模式。前者是指由政府部门(及其附属机构)资助科学数据的生产、采集、存储和管理，并面向社会全体成员提供无偿使用和分享的科学数据，任何组织和个人都可以不受版权协议的限制，对科学数据进行复制、传播和开发。该模式是目前最常用的方式，但也面临后续资助难以为继的困境。后者是指科学数据的资助者与使用者达成许可协议，实行有偿开发利用的一种科学数据共享方式[12]。

按照科学数据来源或规模来看，科学数据开放共享模式包括[52]：①大科学装置的单源数据共享模式，它根据有关研究任务的数据采集计划产生原始数据，存储到服务于该大科学装置的数据中心或服务系统中，优先在参与相应研究任务的合作者范围内共享，一定时间后将部分或全部数据对公众开放。该模式的瓶颈在于数据开放共享的驱动力和资助机构的数据开放政策。②广域合作的分布式注册模式，它按照统一的标准和规范将松散分布的数据、元数据、产品和服务进行集中注册，由统一的数据门户负责连接各注册数据资源，为用户提供一站式数据检索和定位服务，并由分布式的数据服务系统提供数据获取服务。该模式的瓶颈是各子中心开放共享标准和服务可控性问题。③基于数据知识库的集中存缴模式，它利用机构存储库或依托第三方存储库进行数据集中存储和发布。该模式的瓶颈是对存储设施、网络设施、安全设施、管理及发布服务等软硬件环境有极高的要求。④科学数据出版模式，它倡导以出版形式实现数据开放，并通过出版实现数据质量控制以及提供数据再利用的便利渠道。该模式的瓶颈在于确定学术评价体系和数据知识产权等比较困难。⑤数据集市模式，它基于数据中心或存储库的支持，借助数据交易平台或由单一主体提供数据源来实现面向最终用户的交互分享。该模式的瓶颈在于经济收益的可持续性以及数据定价等问题。上述模式描绘了现有科研数据开放共享实践的多种运作方式，但不能把它们与科学数据开放获取途径等同起来，后者往往是通过开放获取数据知识库、开放获取数据期刊和开放获取学术期刊的数据说明文件来实现的[54]。不过，用户免费获取科学数据方式[55]与科研数据开放存取途径并非同一个概念。

此外，如今已经涌现了包括政府部门资助、联盟成员资助、内置成本资助、公私伙伴资助、慈善资助、研究资助、用户资助、服务资助、自愿者资助、众筹、社会支付、虚拟货币支付等在内的多种资助模式[56-57]，其中以政府部门资助模式为主要方式。

5 对科学数据开放共享利益相关者角色与责任的研究

科学数据开放共享的利益相关者包括政府、大学和研究机构、研究资助机构、学术团体和专业协会、出版商、图书馆、研究人员、数据专业人员等，他们可充当科学数据的生产者、资助者、组织者、出版者、传播者、管理者、利用者等多种角色。这些利益相关者中的每一方往往都与特定的开放获取领域、开放获取进程以及特定的数据传播和保存举措相关[41]。

不同利益相关者在科学数据开放共享中可以扮演不同角色，且拥有不同责任。其中，政府负责制定研究系统运作的法律和监管框架，并提供研究资金；研究资助机构负责制定和执行国家研究战略和资助政策，将资金分配给各种研究活动；研究机构应根据学术规范和资助指南，管理其雇员公共资助研究项目的实施，制定访问和使用研究数据的政策，分析研究人员对数据管理的需求，确保每一名研究人员接受适当的研究培训；学术团体和专业协会负责制定和宣传与科学数据开放共享相关的科学规范、价值观和标准；研究人员应参与制定数据注释、保存和长期访问的政策和标准，遵守专业领域的专业标准，在每个研究项目开始时建立数据管理计划，及时公布其公开报告结果所包含的研究数据、方法和其他信息，以便能够核实已发表的结果，除非有令人信服的理由不公布数据，但此时研究人员应该公开解释为什么这些数据会被拒绝发布；数据专业人员应负责数据分析、存档、保存和分发，并参与制定数据标准、格式、元数据和数据质量控制机制[35,58]。此外，研究机构和研究资助者应与研究人员和数据专业人员合作，制定、支持和实施满足研究人员数据需求的计划；研究机构、研究资助机构、专业协会和期刊还应通过出版政策、公开承认和资助等手段促进研究数据的共享[58]7-9。

6 对科学数据开放共享应用与调查的研究

科学数据开放共享应用领域十分广泛,实践调查能够进一步反映科学数据开放共享现状。

6.1 对科学数据开放共享应用的研究

开放科学数据已经在许多领域得到推广和利用,不仅可以用来促进科学研究与创新[59],支持验证临床实验数据结果[60],而且可以用来执行生物保护行动[61]或进行生物多样性保护研究[62]以及远洋生物区划规[63]和罗斯海(the Ross Sea)研究[64],还能够转变教育研究模式[65],提升疟疾研究与控制[66],实现人类基因组项目效用和公共利益的最大化[67]等。在这些应用实践中,开放共享模式并不违反知识产权法具体规定,对于已获得知识产权保护的数据和相关成果,往往通过签订许可合同方式实现开放共享;开放共享模式亦不排斥知识产权保护,即使研究机构或研究者对其获得的数据和相关成果选择进行开放共享,也不意味着完全放弃寻求知识产权保护[68]。不过,仅仅执行强制性数据披露政策可能会减少福利,除非同时有适当的奖励措施来防止拖延[69]。

6.2 对科学数据开放共享调查的分析

从科学数据开放共享调查来看,国内开放科学数据资源主要来源于政府部门、研究型高等院校、网络科学数据库、国际性数据中心,各自分别形成了以基础科学数据为主、以学科专业数据为主且侧重于自然学科、以互联网数据为主、以专有性数据和部分科学数据开放为主的特点[70]。芬兰大学、政府机构、研究机构和市政当局大概保留了数万个数字研究数据集,但只有一小部分被存档和重复使用[49]14,其他国家也存在类似情况。中低收入国家的科学家对数据共享的兴趣不断增长,已经认识到开放科学未来的巨大潜力[71]。

2014 年贝尔蒙特论坛(The Belmont Forum)开放数据调查结果显示,开放数据的主要愿望源于内在的研究动机,如加速科学研究和应用,传播和承认研究成果,个人致力于开放数据,以及数据使用者的要求等;开放科学数据的几个最重要障碍主要是希望在发布数据之前发布结果、法律限制、丧失信用或认可、可能的误解或误用[72]。范·埃因顿(Van den Eynden)等对惠康基金会获得者和 ESRC 资助获得者就"开放研究"的利益与障碍也做了在线调查,发现 80% 的惠康基金获得者和 66% 的 ESRC 资助获得者分别实现了数据集的开放获取[46]。

最近国外学者调查发现,科学家对数据共享和重用持积极态度。科学家们承认,分享科学数据可以在节省时间和提高研究效率方面对科学进步产生积极影响,但在分享自己的研究数据时,科学家们会感到担忧,包括担心这些数据被误用或被曲解。如果能够确保已使用的数据得到适当承认和引用,并让原始研究人员对数据使用有一定程度的控制,那么这些关切就可以得到缓解。另外,帮助研究人员创建元数据、向数据集添加出处信息、寻找足够的公共数据存储库,也可以帮助研究人员采用更好的数据共享做法[73]。

7 对科学数据开放共享对策与建议的研究

为促进科学数据开放共享,现有对策与建议主要包括:①建立健全科学数据开放共享政策法规[74],完善数据共享法规的顶层设计[12];②制定通用的科学数据管理规范与技术标准,提升开放共享各环节关键技术,搭建统一服务平台门户[74];③构建开放共享的科研数据管理制度[75],建立科学数据的公开机制、市场调节机制、合作机制与保障机制[50,76],探索公益性与市场化相结合的道路[74];④建立科学数据质量控制和评估体系[12];⑤建设科学数据开放共享文化[77],普及开放科学的社会效益,不仅为公民就其关心的科学问题得出自己的结论提供所需的证据,而且解释具体工作或生活领域进行数据共享的直接好处[78];⑥加强科学数据管理教育和培训[12];⑦加强开放共享过程中的知识产权保护[74];⑧利益相关者(包括政府、大学和研究机构、学术团体、学术机构和专业机构、慈善机构、研究人员等)发挥其在科学数据开放共享中的作用,积极促进科学数据开放共享的实施[33]71-78;⑨规范科学数据中心运行管理,加快建设一批有重要影响力的国家科学数据中心[79]。

8 结语

综上所述,国内外已经对科学数据开放共享主题进行了广泛研究,涌现了一批研究成果。但是,目前国内外科学数据开放共享理论与实践处于起步阶段,仍有许多不足之处。

首先,从科学数据开放共享理论研究来看,一些基本概念和核心问题没有诠释清楚,这包括:①没有有效区分科学数据开放共享与科学数据共享的差异,有人把科学数据开放共享政策与科学数据共享政策或政府数据开放政策、机构数据政策混为一谈,而事实上这些

概念有本质的区别。今后需要有效界定与区分这几个概念。②没有全面深入阐述科学数据开放共享机制问题,尽管已有学者初步探索该问题(见4.1小节),但是缺少对科学数据开放共享的动力机制、治理机制、协作机制的创新研究,也仍需深化科学数据开放共享保障机制与评价机制研究,以建立健全科学数据开放共享机制。③缺少对我国科学数据开放共享实践与问题的调查分析。因此,需要进一步加强科学数据开放共享理论研究。

其次,从国内科学数据开放共享实践来看,我国科学数据共享服务虽已初见成效,但是国内科学数据政策宏观体系尚待拓展,政策落实仍需继续积累,仍有很多学科门类需要提高数据管理意识和管理水平[80],何况没有形成面向社会开放共享的局面,国家各类科技计划项目所产生的科学数据也没能得到有效的管理和利用[81]。我国与美国、英国等西方发达国家相比,在科学数据开放共享水平上存在明显差距。不过,我国已经认识到科学数据共享的战略意义。《中共中央关于制定国民经济和社会发展第十三个五年规划的建议》明确要求"实施国家大数据战略,推进数据资源开放共享"。国务院办公厅2018年3月17日颁布的《科学数据管理办法》确立了"开放为常态、不开放为例外"的共享原则,而且首次明确了国务院科学技术行政部门、国务院相关部门、省级人民政府相关部门、科研院所、高等院校和企业等法人单位、科学数据中心在科学数据管理(包括开放共享)的职责[82]。因此,今后我国应基于《科学数据管理办法》,逐渐完善科学数据开放共享制度或政策,不断提升我国科学数据开放共享实践水平。

参考文献:

[1] NICOL A, CARUSO J, ARCHAMBAUIT É. Open data access policies and strategies in the European research area and beyond[EB/OL]. [2019－01－15]. http://science-metrix.com/pdf/SM_EC_OA_Data.pdf.

[2] MAUTHNER N S, PARRY O. Open access digital data sharing: principles, policies and practices[J]. Social epistemology: a journal of knowledge, culture and policy, 2013,27(1):47－67.

[3] HOSSAIN M A, DWIVEDI Y K, RANA N P. State-of-the-art in open data research: insights from existing literature and a research agenda[J]. Journal of organizational computing and electronic commerce,2016, 26(1－2):14－40.

[4] 黄永文,张建勇,黄金霞,等. 国外开放科学数据研究综述[J]. 现代图书情报技术,2013(5):21－27.

[5] RUUSALEPP R. A comparative study of international approaches to enabling the sharing of research data[EB/OL]. [2019－01－15]. https://www.era.lib.ed.ac.uk/handle/1842/3361.

[6] 邱春艳,黄如花. 近3年国际科学数据共享领域新进展[J]. 图书情报工作,2016,60(3):6－14.

[7] 扆铁梅,顾立平,董洁,等. 国外科学数据开放获取研究[M]. 北京:中国财政经济出版社,2017.

[8] 张玉娥,王永珍. 欧盟科研数据管理与开放获取政策及其启示——以"欧盟地平线2020"计划为例[J]. 图书情报工作,2017,61(13):70－76.

[9] SPARC EUROPE. An analysis of open data and open science policies in Europe[EB/OL]. [2019－01－15]. http://sparceurope.org/download/1654/.

[10] 张晓青,盛小平. 国外科学数据开放共享政策述评[J]. 图书馆论坛,2018,38(8):147－152.

[11] 崔雁. 科学数据开放共享中的科研利益相关者政策研究[D]. 北京:中国科学院文献情报中心,2016.

[12] 王晴. 论科学数据开放共享的运行模式、保障机制及优化策略[J]. 国家图书馆学刊,2014,23(1):3－9.

[13] 刘细文,熊瑞. 国外科学数据开放获取政策特点分析[J]. 情报理论与实践,2009,32(9):5－9,18.

[14] 尤霞光,盛小平. 8个国际组织科学数据开放共享政策的比较与特征分析[J]. 情报理论与实践,2017, 40(12):40－45.

[15] 温芳芳. 国外科学数据开放共享政策研究[J]. 图书馆学研究,2017 (9):91－101.

[16] 顾立平. 科学数据开放获取的政策研究[M]. 北京:科学技术文献出版社,2016.

[17] SWEENEY L, CROSAS M. An open science platform for the next generation of data[EB/OL]. [2019－01－15]. https://arxiv.org/ftp/arxiv/papers/1506/1506.05632.pdf.

[18] HAEUSERMANN T, GRESHAKE B, BLASIMME A, et al. Open sharing of genomic data: who does it and why? [EB/OL]. [2018－10－08]. https://doi.org/10.1371/journal.pone.0177158.

[19] NAGASAKA Y, SHIMODA K, FUJII N. Multidimensional recording (MDR) and data sharing: an ecological open research and educational platform for neuroscience[EB/OL]. [2019－01－14]. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3141074/.

[20] CHALK S J. The open spectral database: an open platform for sharing and searching spectral data[EB/OL]. [2019－01－15]. https://jcheminf.springeropen.com/articles/10.1186/s13321-016-0170-2.

[21] PEER L, GREEN A. Building an open data repository for a specialized research community: process, challenges and lessons[J]. The international journal of digital curation,2012,7(1):151－162.

[22] BURLEY S, BERMAN H M, CHRISTIE C, et al. RCSB protein data bank: sustaining aliving digital data resource that enables breakthroughs in scientific research andbiomedical education[J]. Protein science,2018, 27(1):316－330.

[23] 朱玲,聂华,崔海媛,等. 北京大学开放研究数据平台建设:探索与实践[J]. 图书情报工作,2016,60(4):44-51.
[24] ABBÀ S, BIRELLO G, VALLINO M,et al. Shall we share? A repository for open research data in agricultureand environmental sciences[J]. Bulletin OEPP/EPPO bulletin,2015,45 (2):311-316.
[25] 罗鹏程,朱玲,崔海媛,等. 基于 Dataverse 的北京大学开放研究数据平台建设[J]. 图书情报工作, 2016,60(3): 52-58.
[26] MÜLLER J, HEISSK I, OBERHOFFER R. Implementation of an open adoptionresearch data management system for clinical studies [EB/OL]. [2019-01-15]. https://link.springer.com/article/10.1186/s13104-017-2566-0.
[27] BHAT M H. Exploring research data in Indian institutional repositories[J]. Program: electronic library and information systems, 2014,48(2):206-216.
[28] 邱春艳. 欧盟科学数据开放存取实践及启示[J]. 情报理论与实践,2016,39(11):138-144.
[29] TSIMPOGOU F, KOUOUNIDOU V V, PROKOPIOU L. Open access to scientific results and data. European Union's efforts through OpenAIRE and OpenAIREplus FP7 projects: cypriotparticipation[EB/OL]. [2019-01-15]. https://core.ac.uk/download/pdf/38296431.pdf.
[30] RETTBERG N, SCHMIDT B. OpenAIRE-building a collaborative open access infrastructure for European researchers[J]. Liberquarterly,2012,22(3):160-175
[31] LUZI D, RUGGIERI R, PISACANE L. The OpenUP pilot on research data sharing, validation and dissemination in social sciences [C]//Digital libraries: supporting open science. Basel: Springer International Publishing,2019:248-258.
[32] STEPHEN E F, MARGARET E M, MIRON L S. Sharing research data [M]. Washington, DC: National Academy Press, 1985.
[33] THE ROYAL SOCIETY. Science as an open enterprise[M]. London:The Royal Society,2012.
[34] JOSEPH H. The evolving U. S. policy environment for open research data[J]. Information services & use,2016,36 (1-2): 45-48.
[35] UHLIR P F, SCHRÖDER P. Open data for global science [EB/OL]. [2019-01-15]. https://datascience.codata.org/articles/10.2481/dsj.6.OD36/galley/367/download/.
[36] STAGARS M. Open data in Southeast Asia[M]. Singapore: Palgrave Macmillan,2016.
[37] HOUGTON J, SHEEHAN P. Estimating the potential impacts of open access to research findings[J]. Economic analysis and policy, 2009,29(1):127-142.
[38] JANSSEN M, CHARALABIDIS Y, ZUIDERWIJK A. Benefits, adoption barriers and myths of open data and open government[J]. Information systems management,2012,29(4):258-268.
[39] KOWALCZYK S,SHANKAR K. Data sharing in the sciences[J]. Annual review of information science and technology, 2010,45 (1):247-294.
[40] BORGMAN C L. The conundrum of sharing research data[J]. Journal of the American Society for Information and Technology, 2012, 63(6):1059-1078.
[41] WESSELS B, FINN R L, LINDE P, et al. Issues in the development of open access to research data[J]. Prometheus, 2014,32 (1):49-66.
[42] FRY J, LOCKYER S, OPPENHEIM C, et al. Identifying benefits arising from the curation and open sharing of research data produced by UK Higher Education and Research Institutes[EB/OL]. [2019-01-15]. https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/4600/1/JISC%20data%20sharing_final%20report.pdf.
[43] ONYANCHA O B. Open research data in Sub-Saharan Africa: a bibliometric study using the data citation index[J]. Publishing researchquarterly,2016,32(3): 227-246.
[44] 姚乐野. 加快社会科学数据资源的开放共享,促进中国特色哲学社会科学的构建[J]. 西华大学学报(哲学社会科学版), 2018,37(3):68-72.
[45] PIWOWAR H, DAY R S, FRISMA D B. Sharing detailed research data is associated with increased citation rate [EB/OL]. [2019-01-15]. http://journals.plos.org/plosone/article? id = 10.1371/journal.pone.0000308.
[46] VAD DEN EYNDEN V, KNIGHT G, VLAD A, et al. Towards open research: practices, experiences, barriers and opportunities [EB/OL]. [2019-01-15]. https://ndownloader.figshare.com/files/6739038.
[47] 郭仕琳. 政策视角下科学数据开放共享的利益平衡机制研究[D]. 哈尔滨:黑龙江大学,2018.
[48] SCHOPFEL J, CHAUDIRON S, JACQUEMIN B, et al. Open access to research data in electronic theses and dissertations: an overview[J]. Libraryhi tech,2014,32(4):612-627.
[49] KUULA A, BORG S. Open access to and reuse of research data - the state of the art in Finland[M]. Tampere: Finnish Social Science Data Archive,2008.
[50] 陈传夫. 中国科学数据公共获取机制:特点、障碍与优化的建议[J]. 中国软科学,2004,19(2):8-13.
[51] 万望辉,崔辰州,乔翠兰,等. 天文观测数据开放共享政策与策略分析研究[J]. 天文研究与技术,2015,12(3): 364-373.
[52] 诸云强,朱琦,冯卓,等. 科学大数据开放共享机制研究及其对环境信息共享的启示[J]. 中国环境管理, 2015,7(6):38-45.
[53] 李成赞,张丽丽,侯艳飞,等. 科学大数据开放共享:模式与机制[J]. 情报理论与实践,2017,40(11):45-51.
[54] 刘晶晶,马建华. 论科研数据开放共享的三种途径[J]. 情报杂志,2015,34(10):146-150,96.

[55] 黄如花,刘龙,陈萌. 科学数据开放存取的途径[J]. 图书馆,2016(11):26－33.

[56] KITCHIN R, COLLINS S, FROST D. Funding models for open access digital data repositories [J]. Online information review, 2015, 39(5):664－681.

[57] EISFELD-RESCHKE J, HERB U, WENZLAFF K. Research funding in open science[M]//BARTLING S, FRIESIKE S. Opening science: the evolving guide on how the Internet is changing research, collaboration and scholarly publishing. Heidelberg:Springer,2014:237－253.

[58] COMMITTEE ON ENSURING THE UTILITY AND INTEGRITY OF RESEARCH DATA IN A DIGITAL AGE. Ensuring the integrity, accessibility, and stewardship of research data in the digital age [M]. Washington, DC: The National Academies Press,2009.

[59] XU G H. Open access to scientific data: promoting science and innovation[EB/OL]. [2019－01－15]. https://datascience.codata.org/articles/abstract/10.2481/dsj.6.OD21/.

[60] PENCINA M J, LOUZAO D M, McCOURT B J, et al. Supporting open access to clinical trial data for researchers: the Duke Clinical Research Institute-Bristol-Myers Squibb supporting open access to researchers initiative[J]. American heart journal,2016,172:64－69.

[61] SULLIVAN B L, PHILLIPS T, DAYER A A, et al. Using open access observational data for conservation action: a case study for birds[J]. Biological conservation, 2017,208:5－14.

[62] TURNER W, RONDININI C, PETTORELLI N, et al. Free and open-access satellite data are key to biodiversity conservation[J]. Biological conservation,2015,182:173－176.

[63] ROBERSON L A, LAGABRIELLE E, LOMBARD A T, et al. Pelagic bioregionalisation using open-access data for better planning of marine protected area networks[J]. Ocean & coastal management, 2017,148:214－230.

[64] HUETTMANN F, SCHMID M S, HUMPHRIES G R W. A first overview of open access digital data for the Ross Sea: complexities, ethics, and management opportunities[J]. Hydrobiologia, 2015, 761(1):97－119.

[65] GILMORE R O, ADOLPH K E, MILLMAN D S, et al. Transforming education research through open video data sharing [EB/OL]. [2019－01－15]. https://files.eric.ed.gov/fulltext/EJ1106045.pdf.

[66] MOYES C L, TEMPERLEY W H, HENRY A J, et al. Providing open access data online to advance malaria research and control [EB/OL]. [2019－01－15]. https://malariajournal.biomedcentral.com/articles/10.1186/1475－2875－12－161.

[67] PEREIRA S, GIBBS R A, McGUIRE A L. Open access data sharing in genomic research[J]. Genes,2014, 5(3):739－747.

[68] 伍春艳,焦洪涛,范建得. 人类遗传数据的开放共享抑或知识产权保护[J]. 知识产权,2014(1): 55－60.

[69] MUELLER-LANGER F, ANDREOLI-VERSBACH P. Open access to research data: strategic delay and the ambiguous welfare effects of mandatory data disclosure[J]. Information economics and policy,2018,42:20－34.

[70] 李志芳,邓仲华. 国内开放科学数据的分布及其特点分析[J]. 情报科学,2015,33(3):45－49.

[71] BEZUIDENHOUT L, CHAKAUYA E. Hidden concerns of sharing research data by low/middleincomecountry scientists[J]. Journal of empirical research on human research ethics, 2015,10(3): 302－313.

[72] SCHMIDT B, GEMEINHOLZER B, TRELOAR A. Open data in global environmental research: the Belmont Forum's open data survey[EB/OL]. [2019－01－15]. http://journals.plos.org/plosone/article? id=10.1371/journal.pone.0146695.

[73] TENOPIR C, CHRISTIAN L, ALLARD S, et al. Research data sharing: practices and attitudes of geophysicists [J]. Earth and space science, 2018,5:891－902.

[74] 马宁,刘召. 科学数据资源开放共享体系研究[J]. 中国科技资源导刊,2017,49(3):1－7.

[75] 于风程. 构建开放共享的科研数据管理制度[N]. 学习时报,2018－03－02(6).

[76] 是沁,储节旺. 基于系统动力学的科学数据开放共享保障机制研究[J]. 情报杂志,2018,37(11):143－149.

[77] PAMPEL H, DALLMEIER-TIESSEN S. Open research data: from vision to practice[M]//BARTLING S, FRIESIKE S. Opening science: the evolving guide on how the Internet is changing research, collaboration and scholarly publishing. Heidelberg:Springer,2014:213－224.

[78] BISOL G D, ANAGNOSTOU P, CAPOCASA M, et al. Perspectives on open science and scientific data sharing: an interdisciplinary workshop[J]. Journal of anthropological sciences,2014,92:179－200.

[79] 袁于飞. 让科学数据开放共享成为常态[N]. 光明日报,2018－04－05(2).

[80] 张丽丽,温亮明,石蕾,等. 国内外科学数据管理与开放共享的最新进展[J]. 中国科学院院刊, 2018, 33(8): 774－782.

[81] 杨兰. 开放科学数据,实现共同发展——中国科学数据共享之路[N]. 科技日报,2014－04－24(12).

[82] 国务院办公厅. 国务院办公厅关于印发科学数据管理办法的通知[EB/OL]. [2019－01－15]. http://www.most.gov.cn/mostinfo/xinxifenlei/fgzc/gfxwj/gfxwj2018/201804/t20180404_139023.htm.

作者贡献说明:

盛小平:论文撰写与修改;

武彤:资料收集与分析。

Review on Open Sharing of Scientific Data Across the World

Sheng Xiaoping[1] Wu Tong[2]
[1] School of Library, Information and Archives, Shanghai University, Shanghai 200444
[2] School of Economics and Management, South China Normal University, Guangzhou 510006

Abstract: [**Purpose/significance**] By summarizing the research results of open sharing of scientific data across the world, this paper provides a reference for promoting the development of relevant theoretical research and practice in China. [**Method/process**] By using text content analysis method, this paper analyzes the research progress across the world from seven aspects: open sharing policies of scientific data, platforms and projects, benefits and obstacles, mechanisms and models, stakeholder roles and responsibilities, applications and investigations, andstrategies and suggestions. [**Result/conclusion**] At present, the topic of open sharing of scientific data has been extensively explored across the world, especially on policies and platforms of open sharing of scientific data, mechanisms and modes, benefits and obstacles, and so on. But on the whole, it is still in its infancy, so it is necessary to improve the research level of open sharing of scientific data from the aspects of distinguishing between the concept and behavior of open sharing of scientific data and sharing of scientific data effectively, establishing and perfecting the mechanism of open sharing of scientific data and strengthening the investigation and analysis of the practice of open sharing of scientific data.

Keywords: scientific data open sharing research review